

Web tools to analyze omic data

**MDA course on NGS Data Analysis
Valencia, 28 Sep 2015**



PRINCIPE FELIPE
CENTRO DE INVESTIGACION

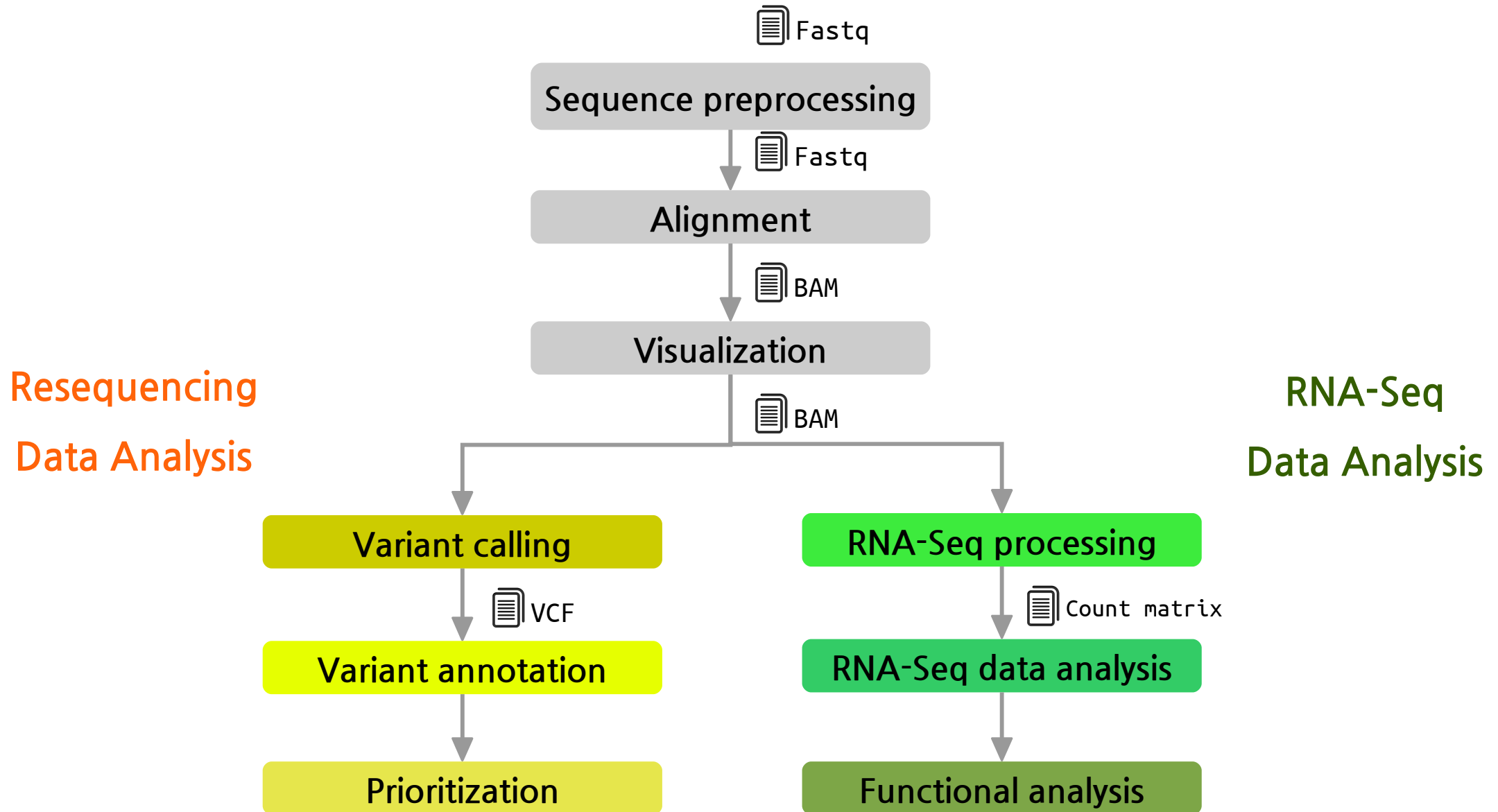
Computational · Genomics



Francisco García
fgarcia@cipf.es

MDA course on NGS Data Analysis

NGS Data Analysis Pipeline



Genome Maps

The screenshot displays the Genome Maps web browser interface for Homo sapiens GRCh37.p10. The main view shows a genomic region with coordinates 13:32889611-32973805. The interface is divided into several sections:

- Region overview:** Shows a window size of 673,559 nts. It displays various genomic features including lincRNAs (e.g., <FRY-AS1 [antisense]), protein-coding genes (e.g., FRY, ZAR1L, BRCA2, N4BP2L2, N4BP2L1, RP11-298P3.4), pseudogenes (e.g., ATP8A2P2, FIT1P1), and snoRNAs (e.g., SNORA16).
- Detailed information:** Shows a window size of 84,195 nts. It includes a "Sequence" track with a "Zoom in to view the sequence" button, a "Gene" track, and a "SNP" track.
- Active tracks:** A sidebar on the right lists active tracks: Sequence, Gene/Transcript, SNP, and Mutation, each with a toggle and a close button.
- Navigation and Search:** The top navigation bar includes a search bar with the text "gene, snp, ...", navigation arrows, and a "Configure" button.

Introduction

A next-generation web-based genome browser

CIBERER Spanish Variant Server



Clear Search

Position

Chromosomal Location:

1:1-100000,2:1-100000

Gene:

BEST1

Subpopulations

- MGP (267 healthy controls)
- IBS (107 Spanish individuals from 1000genomes)
- I Certain infectious and parasitic diseases
- II Neoplasms
- III Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
- IV Endocrine, nutritional and metabolic diseases
- V Mental and behavioural disorders
- VI Diseases of the nervous system
- VII Diseases of the eye and adnexa

Chr	Position	Alleles	Id	MAF							1000G MAF (phase 1)					
				Genotypes				Freq.			ALL	AME	ASI	AFR	EUR	ALL
				0/0	0/1	1/1	./.	0 Freq	1 Freq	MAF						
11	61717337	G>C	rs972355	577	1	0	0	0.999	0.001	0.001	0.380	0.470	0.070	0.220	0.320	0.320
11	61717607	C>T	rs972354	575	3	0	0	0.997	0.003	0.003	0.310	0.270	0.280	0.170	0.100	0.340
11	61717715	T>C	rs972353	569	9	0	0	0.992	0.008	0.008	0.380	0.470	0.070	0.220	0.320	0.320
11	61719334	G>A	.	577	1	0	0	0.999	0.001	0.001
11	61719355	G>C	.	577	1	0	0	0.999	0.001	0.001
11	61719387	T>C	rs1800007	340	195	35	8	0.768	0.232	0.232	0.380	0.470	0.070	0.210	0.320	0.310
11	61719458	G>C	rs200543078	577	1	0	0	0.999	0.001	0.001	0.001	0.003	0.000	0.000	0.000	0.000
11	61722627	G>C	rs1801393	560	17	1	0	0.984	0.016	0.016	0.030	0.110	0.020	0.004	0.010	0.030
11	61722645	C>A	rs1109748	512	64	2	0	0.941	0.059	0.059	0.270	0.250	0.290	0.090	0.080	0.310
11	61722811	G>A	.	577	0	1	0	0.998	0.002	0.002

Page 1 of 6 items 1 - 10 of 60

Genomic context Frequencies Phenotype Effect

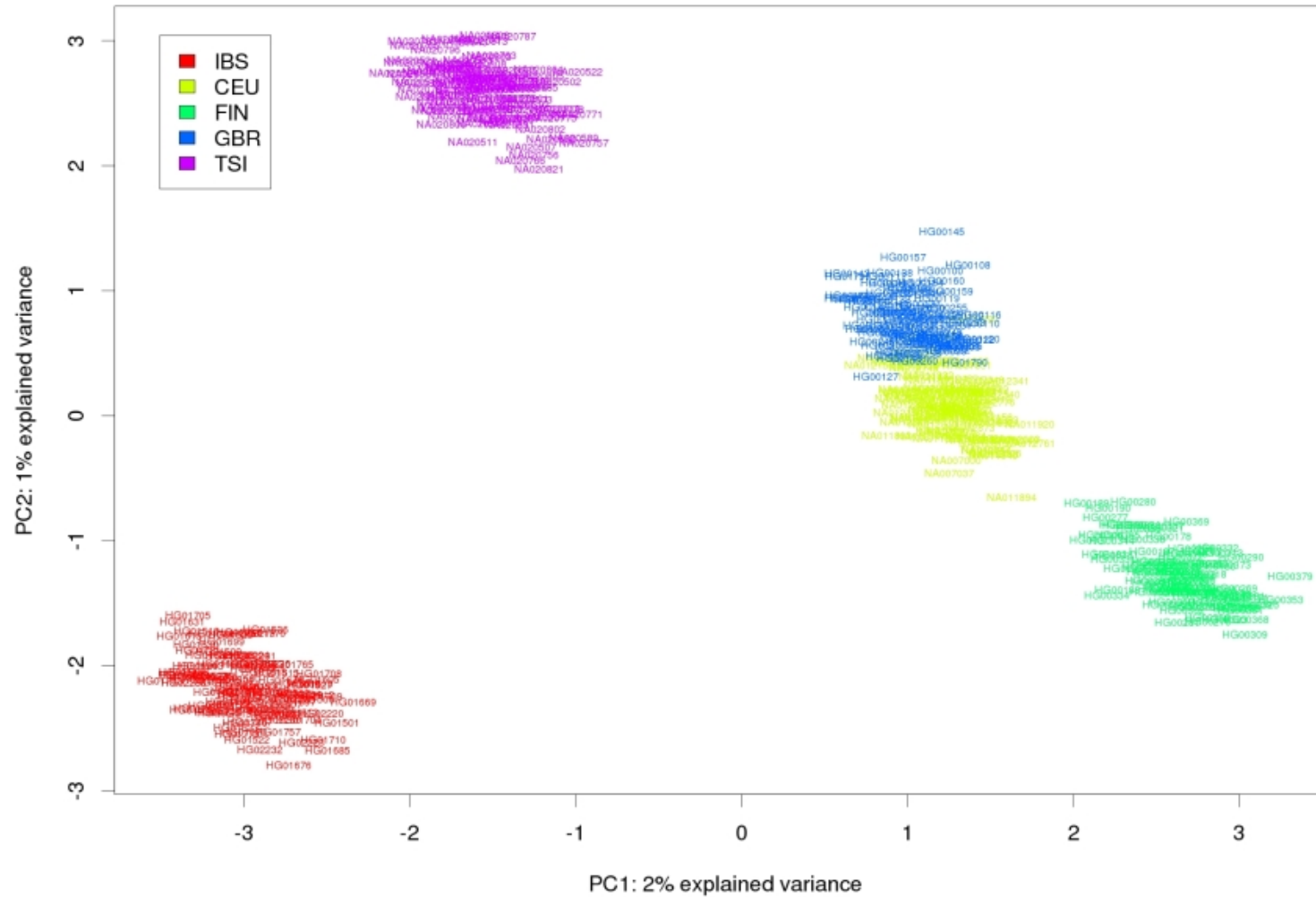
Min - + Max 85 1:1004566-1004650 Go! << < > >>

Region overview Window size: 679 nt

1004269 1004608 1004947

CIBERER Spanish Variant Server

PCA plot for European populations



Introduction

Local genetic variability

BiERapp

Primary
Analysis

1. Sequence preprocessing

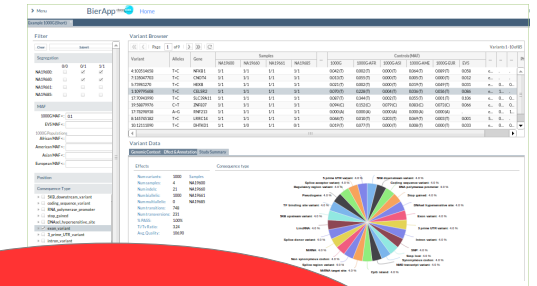
2. Mapping

3. Variant calling

→ **VCF files**

Secondary

4. Variant prioritization

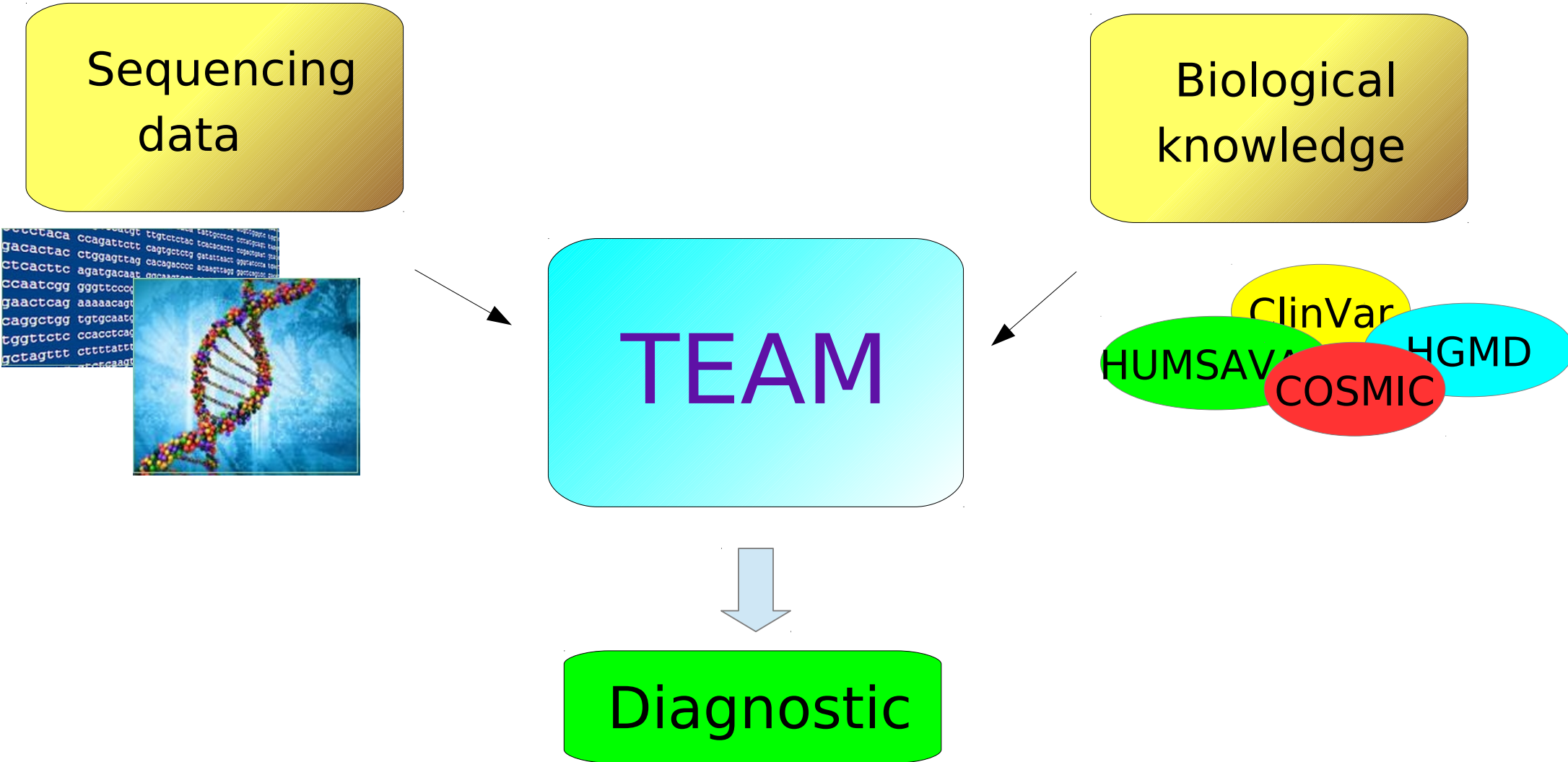


BiERapp

Introduction

BiERapp: discovering variants

Can I interpret sequencing data for diagnostic?



Babelomics 5

Primary

1. Sequence preprocessing

2. Mapping

3. Quantification

Secondary

4. Normalization

5. Differential expression

6. Functional Profiling



Introduction

RNA-Seq Data Analysis